The Atlantic

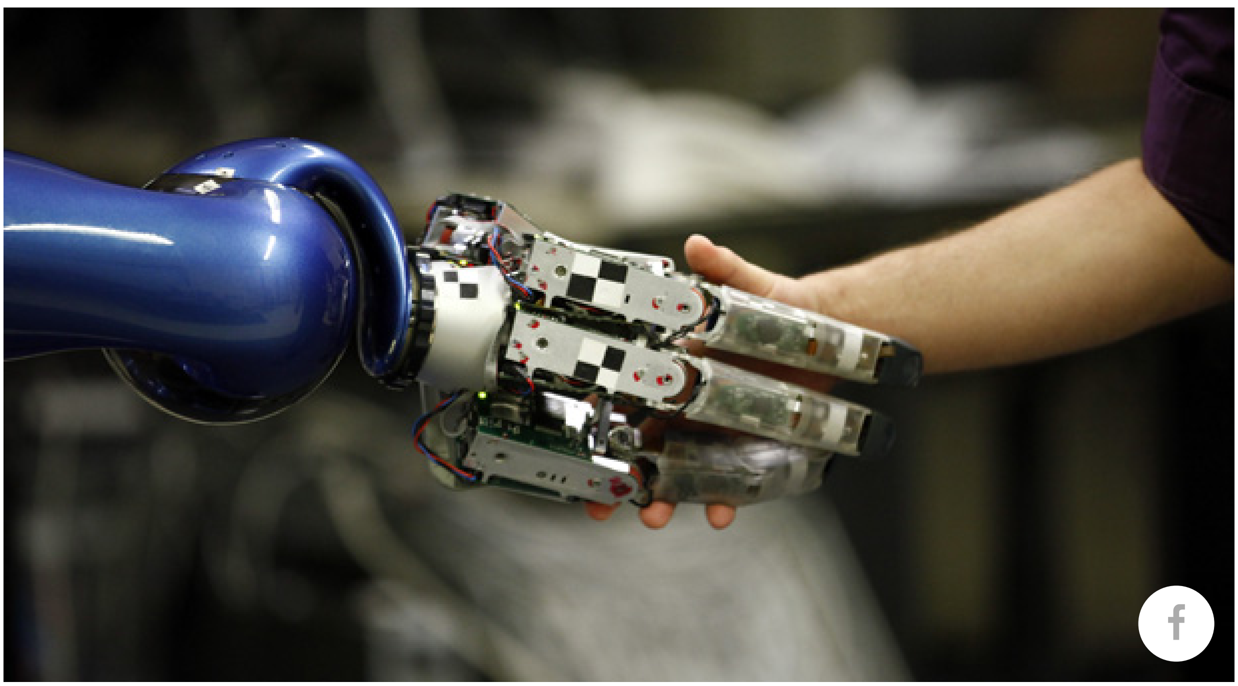
# A Conjecture for a Better Turing Test

BARBARA J. GROSZ

AUG 29, 2012 | TECHNOLOGY



*What sort of philosophical standard should we set for today's powerful computer technology?*



Reuters

Alan Turing is most well known for giving us the universal digital computing machine, the idea that numbers could encode machine instructions as well as be data for those instructions, the vision of computing machines that were more than fast calculators along with some hints on how one might go about addressing the challenge of building a thinking machine, and the Turing Test. Turing devised the Test, which he called called "the imitation game", to replace the philosophical question of whether computing machines could think by an empirical, operational test. It cleverly avoids the need to define intelligence. To succeed in this game, a computer must be able to carry on a dialogue, in natural language, well enough to be indistinguishable from a person. Importantly, there are no constraints on the range of topics to be discussed. Turing asks whether there are "imaginable digital computers which would do well" at this task. This challenge has generated decades of important research and many debates.

> *Is it imaginable that a computer-agent team-member could behave, over the long-term and in uncertain, dynamic environments, in such a way that people on the team will not notice it is not human?*

At the end of the 1950 *Mind* paper (pdf), in which the Test is defined, Turing put forth the very long-term goal of computing machines having competence in "all purely intellectual fields" and notes that opinions vary as to which approach is the best one to follow to achieve this competence. The tension that he described, between building expertise into systems and building systems that learn, remains prevalent. Turing's suggestion that multiple approaches be tried is good advice in general. More strikingly, some of the most visible AI applications, including Watson and Siri, embed this approach in the very way they operate.

Turing ended a BBC interview in 1951 by discussing various ways in which the whole idea of machines thinking is unsettling to some people. Those qualms exist to this day, and there are even more arguments against the possibility now than there were then. Turing's final words in this interview echo still. For me, they express one of the most important reasons for continuing the quest. He says, "The whole thinking process is still rather mysterious to us, but I believe that the attempt to make a thinking machine will help us greatly in finding out how we think ourselves."

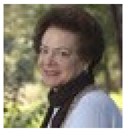
The sentence in the *Mind* paper I like best, though, is this one: "Conjectures are of great importance since they suggest useful lines of research." At a time when so much focus in education, industry and even research is on the short term, on following paths sure to deliver results, it is crucial to look far afield, to conjecture about what might be and to imagine different futures. With this in mind, I have been asking myself and others in computer science what question Turing might pose today were he around to see the vastly increased power of computing machines and the many ways in which computers are deployed now, which are so very different from his day's one-on-one person-computer interactions focused on mathematical computations.

I've made one proposal for this new conjecture, a challenge that reflects the great advances since 1950 in computer science, neuroscience and the behavioral sciences as well as the highly networked ways in which ordinary people now use computers daily: Is it imaginable that a computer-agent team-member could behave, over the long-term and in uncertain, dynamic environments, in such a way that people on the team will not notice it is not human?

Unlike Turing's original question, this question asks not that a computer agent be indistinguishable from a person, but that it behave reasonably, that its mistakes make sense and are not noticeably non-human. It's here that even Watson and Siri reveal their nature: it's in their errors that systems make most evident they are not yet thinking like Turing imagined they might. This question raises the challenge of building systems that work better for people, because they are smarter about matching their abilities with people's. It lets us imagine health care systems that help physicians, nurses, social workers, and patients' families do the work of caring better, rather than imposing on them data entry tasks or taxing interfaces. It suggests a vision of computer agents that support children in exploring biological, physical and mathematical worlds so that learning is fun and integrated with life. It imagines a future in which computer systems make us feel smarter, not dumber, and work seamlessly with us, like a good human partner.



## ABOUT THE AUTHOR



**BARBARA J. GROSZ** is Higgins Professor of Natural Sciences in the School of Engineering and Applied Sciences at Harvard University. She is a member of the National Academy of Engineering, the American Philosophical Society, and the American Academy of Arts and Sciences.